

# Creating a Search Engine for Uniform Titles of Musical Works

Jaska Järvilehto

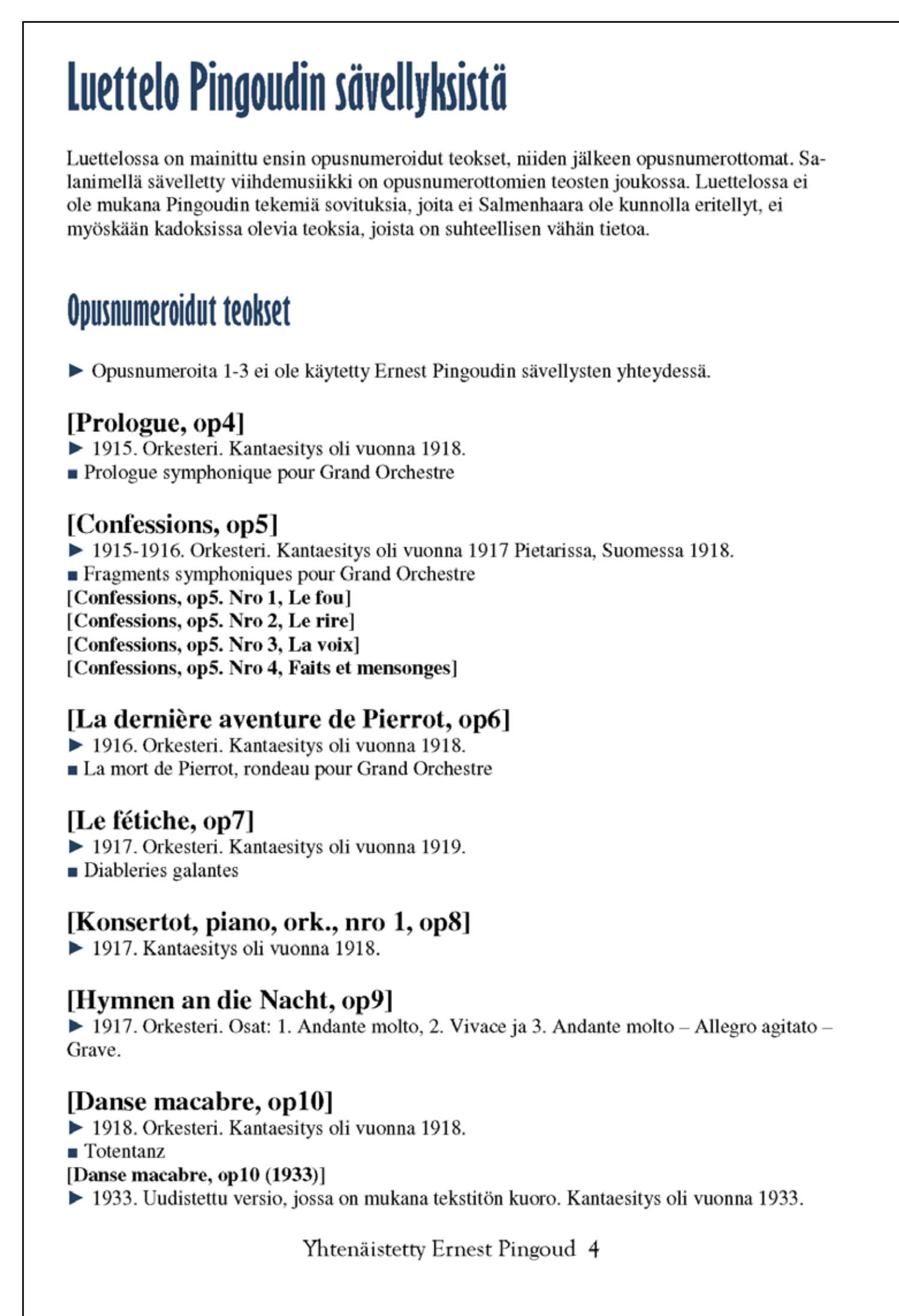
Finnish Music Library Association (IAML Finland)  
jaska.jarvilehto@gmail.com

## Uniform Titles of Musical Works

In the Anglo-American Cataloguing Rules (AACR) and cataloguing rules derived from it, such as the now obsolete Finnish Cataloguing Rules, a uniform title is a title that identifies a specific work in a library catalogue.

Since 1988, the Finnish Music Library Association has published catalogues of uniform titles for the works of 71 composers, first distributed in print form and later as PDF files online<sup>1</sup>, to aid Finnish music cataloguers in the use of properly formed uniform titles.

Finnish Cataloguing Rules were superseded by the Resource Description & Access (RDA) standard<sup>2</sup> in 2016. The Finnish Expert Group on Music Metadata<sup>3</sup>, nominated by the National Library of Finland, has stated that the catalogues of uniform titles are still the primary source of authorized access points for musical works, even though many of the titles in the catalogues are not RDA compliant. Revision of the catalogues, accumulated over several decades, has been postponed ever since.



Example page from the catalogue of uniform titles of the works of Ernest Pingoud by Heikki Poroila<sup>4</sup>.

## Data Extraction and Storage

In 2020 and 2021, a set of tools was created to extract uniform title data from the PDF files, to store and edit the data in a database, and to provide access to the data through a publicly accessible search engine.

A Python script is used to extract uniform titles from a PDF file to a CSV file. The CSV file is uploaded to a web application and the uniform titles, along with any extra information, are stored to a SQL database. The web application enables manual annotation, editing and enriching of the data in the database. Person names, place names, music genres, instruments and mediums of performance are linked to Finto<sup>5</sup>, a centralised service for interoperable thesauruses, ontologies and classification schemes for different subject areas, developed by the National Library of Finland.

## Data Access

The uniform title data in the database is mirrored in real time to an ElasticSearch index. The index is searchable via a custom search engine web application. The search engine<sup>6</sup> currently contains the uniform titles for the works of the Finnish composers Joonas Kokkonen (1921–1996) and Ernest Pingoud (1887–1942). Work is being carried out to include more composers in the database in the future. The goal is to include all 71 composers for whom catalogues of uniform titles currently exist in PDF format. All uniform title data published on the search engine is dedicated to the public domain under the Creative Commons CC0 licence.



Index page for the uniform titles of the works of Ernest Pingoud.

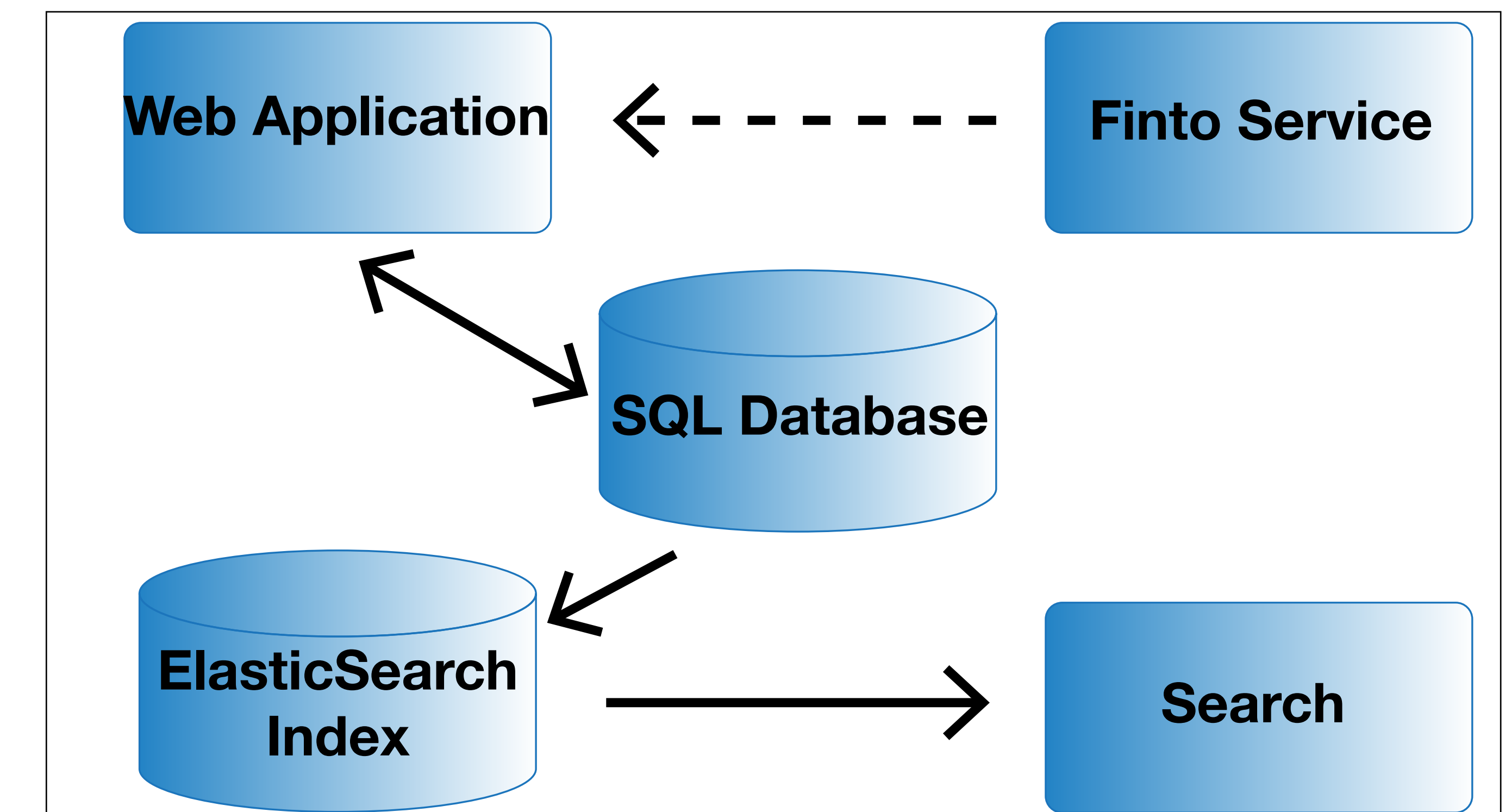


Diagram of the web application and data storage. Arrows indicate the flow of data.

## Current Work and Future Plans

In the beginning of 2022, a private company was founded to maintain the tools and to update the web application, while the Finnish Expert Group on Music Metadata is responsible for the revision and updating of the uniform titles.

There are plans to publish the uniform title data as machine readable data sets on GitHub in the future.

In the optimal future scenario the National Library of Finland would establish a centralized registry for authorized access points for musical works based on the uniform title data.

## References

- <https://www.musiikkikirjastot.fi/tiedonhaunavuksi/saveltajien-teosluetteloita-kirjastojen-tapaan/>
- <https://www.rdatoolkit.org/>
- <https://www.kiwi.fi/display/muusa>
- <https://www.musiikkikirjastot.fi/wp-content/uploads/2016/11/Pingoud.pdf>
- <https://finto.fi/en/>
- <https://musiikinluettelo.fi/ohjeluetelot>